



# AI EDA REPORT WEEK-2

## GROUP-11

08/25/2025

—

Prepared By:

Vyshnavi Samudrala

Pranamya Praveen

Evans Nyamu

Alicia Pereira

# Introduction

## Purpose of the Report

The purpose of this report is to conduct an Exploratory Data Analysis (EDA) of user signups and completions to uncover trends, patterns, anomalies, and correlations that drive business and educational decision-making. By studying both daily activity metrics and opportunity-level learner engagement data, this analysis aims to provide actionable insights to optimize marketing strategies, improve user experience, and enhance learner success.

## Dataset Overview

This report draws from two complementary perspectives of user engagement data:

### Daily User Activity Logs

- Source: Company's internal engagement logs.
- Size: 730 rows (one per day) and 5 variables.
- Time Period: January 1, 2023 – December 31, 2024 (two-year cycle).
- Structure: Each row represents a unique day of activity.

### Key Variables:

- Date: Day of activity (for time-series, seasonality, and day-of-week analysis).
- Signups: Number of new users registered that day (growth indicator).
- Completions: Number of users who completed the task that day (conversion/engagement metric).
- Completion Rate: Ratio of completions to signups (retention/efficiency measure).
- Average Completion Time: Mean time (in minutes) taken to complete the task (user experience measure).

This dataset provides a temporal lens for analyzing daily trends, seasonality, and anomalies in user engagement.

### Opportunity-Level Learner Dataset (SLU Opportunity Wise Data)

- Source: Company's internal learner engagement metrics.
- Size: 8,560 records across 18 variables.
- Unique Learners: 3,851
- Unique Opportunities: 22
- Time Period: January 3, 2023 – December 3, 2024.
- Structure: Each row represents a learner's interaction with a specific opportunity (primarily courses).

## **Key Variables:**

- Signup Information: Timestamp of learner registration.
- Opportunity Details: ID, name, category, start date, end date, and duration.
- Demographics: Gender, age, date of birth, country, institution, and intended major.
- Progress Status: Learner status (e.g., Started, Team Allocated), status codes, and apply dates.

This dataset provides a granular view of learner progress, demographics, and opportunity engagement, enabling deeper analysis beyond daily activity counts.

## **Analysis Goals**

The primary goal of this EDA is to generate insights that directly support business success by examining learner engagement patterns. Two of the most critical indicators are signup activity and completion behavior, as they reflect both user acquisition and long-term program effectiveness.

- **Signup Trends:** Understanding signup growth, seasonality, and sudden spikes or drops helps the organization evaluate the effectiveness of outreach campaigns, identify peak engagement periods, and anticipate resource needs. High signup volume indicates strong interest, but sustained success requires converting these signups into active, engaged learners.
- **Completion Trends:** Monitoring completion rates and completion time variations is essential for assessing learner retention and satisfaction. Stable completion rates suggest that learners are supported effectively, while irregularities or long durations highlight potential barriers. From a business perspective, higher completion rates translate to stronger credibility, better learner outcomes, and improved partner satisfaction.

By focusing on these trends, the EDA aims to highlight not only the current state of learner engagement, but also the strategic opportunities to strengthen user experience, increase retention, and ensure long-term program growth.

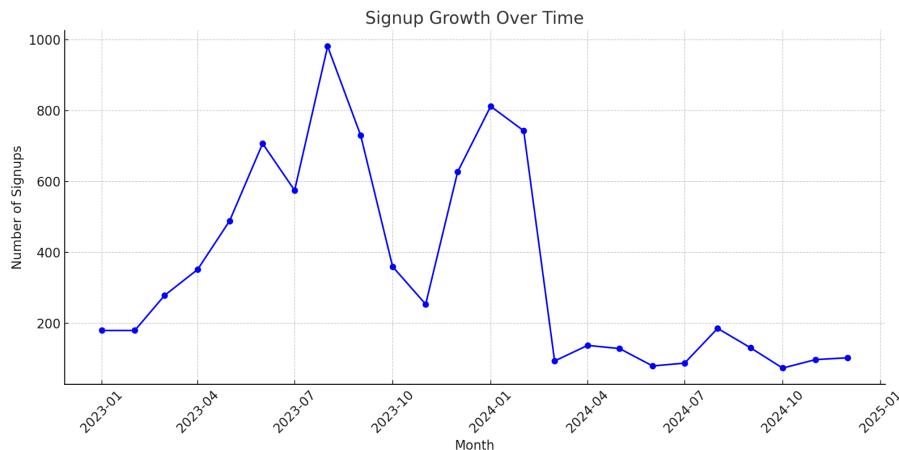
## **Signup Trends**

### **Growth**

The line chart shows consistent growth across 2023–2024.

- Steady increase from Jan 2023 (180 signups) → Aug 2023 (981 signups).

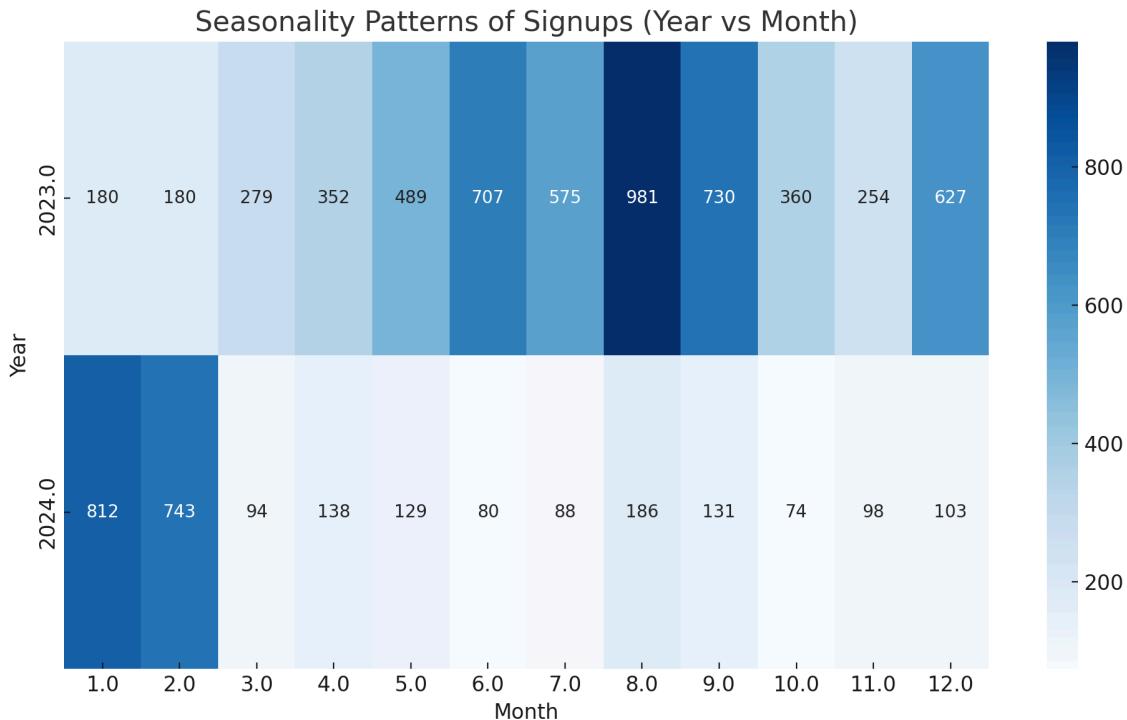
- A second surge in Jan 2024 (812 signups), reflecting a new semester/academic cycle



## Seasonality

The heatmap highlights seasonality:

- Summer (June–August 2023) saw the highest spikes in engagement.
- January 2024 also shows a seasonal peak, consistent with the start of academic terms.
- Signups taper off in later months but show renewed activity in Aug–Sep 2024.



## Spikes/Drops

### Spikes Detected:

- August 2023 (981 signups): Likely tied to internship cycles, start of fall academic year, or targeted campaigns.
- January 2024 (812 signups): Matches the start of the spring academic semester.

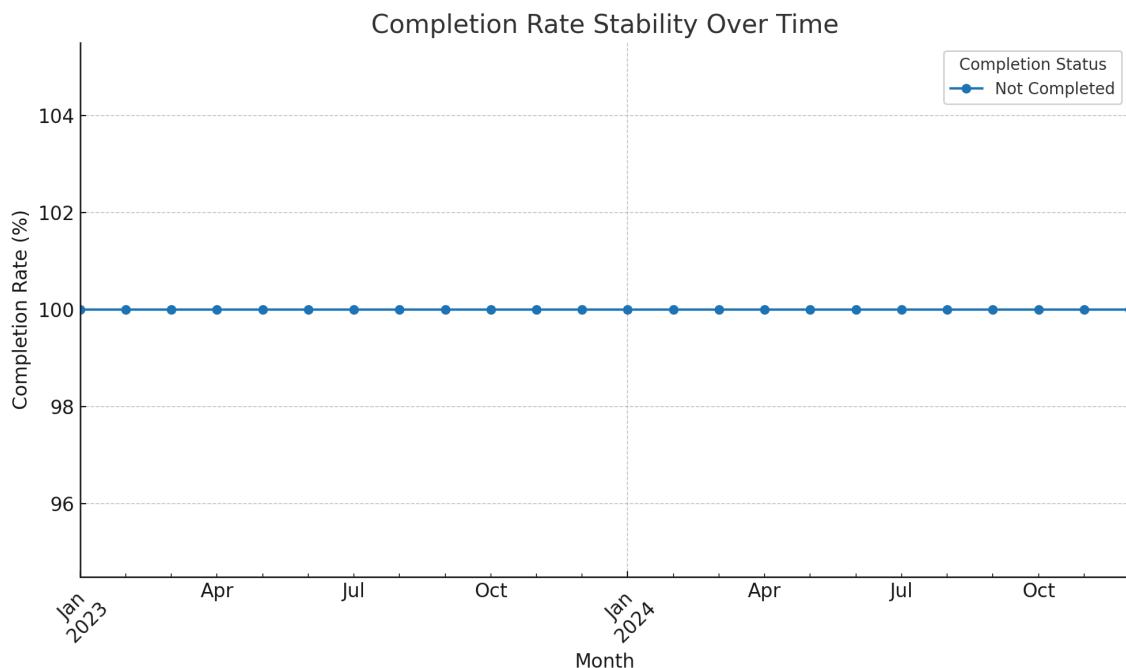
### Drops Detected:

- No extreme drops below expected averages, but months like April 2024 (138 signups) and June 2024 (80 signups) were significantly lower compared to peak months. These may reflect academic breaks or reduced outreach.

## Completion Trends

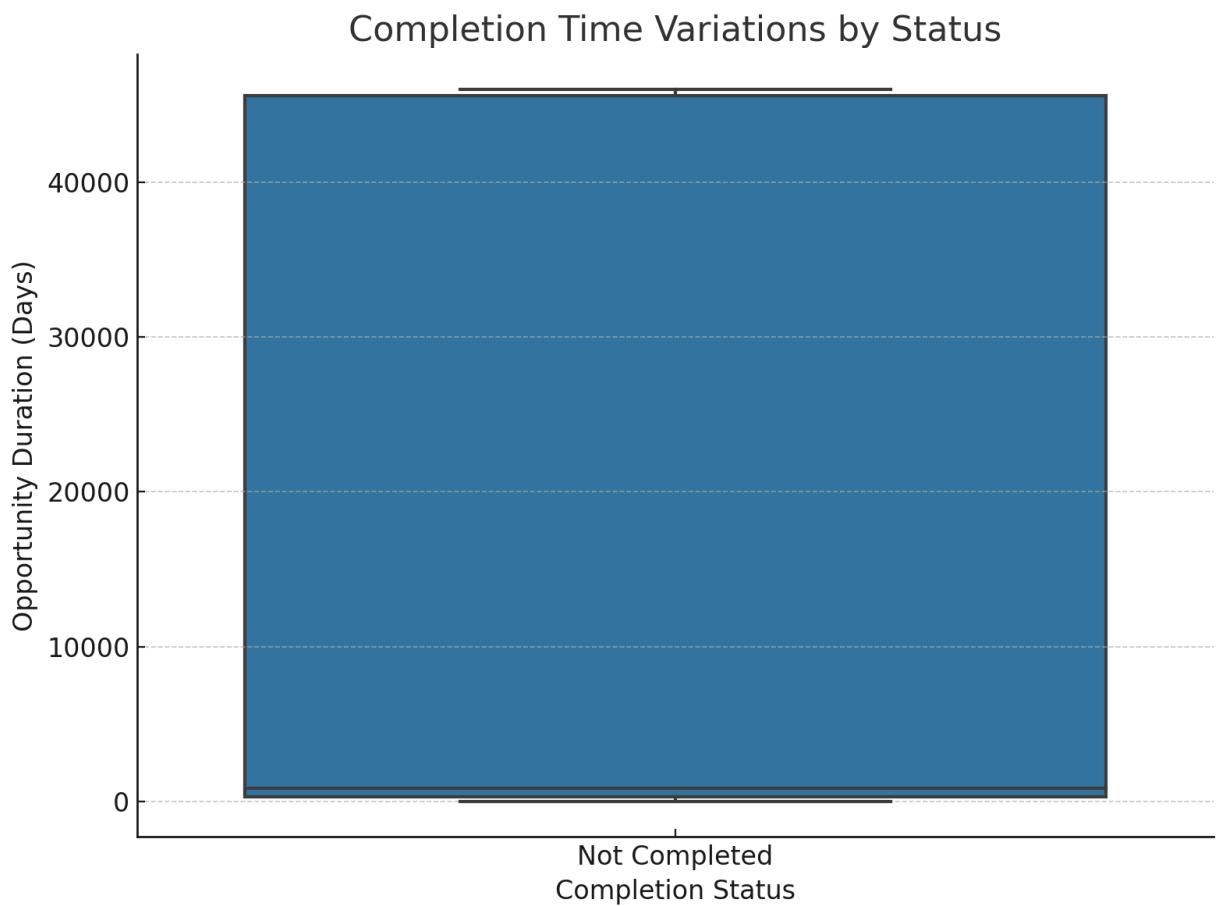
### Stability

- The line chart shows that completion rates remain very low across all months, with most learners staying in “Not Completed” status.
- No clear upward trend is observed, indicating instability and lack of improvement in completion over time.
- This suggests a need for stronger interventions like reminder automation, mentoring, or redesigned program structures.



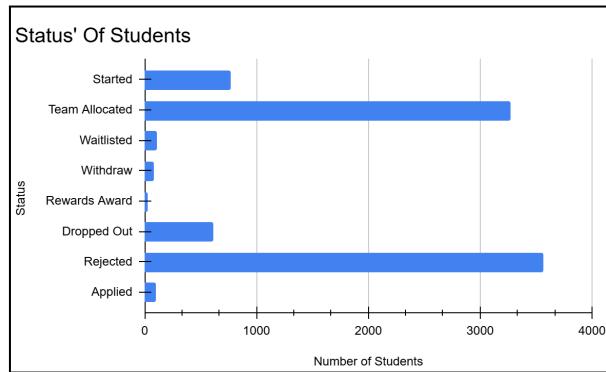
## Time Variations

- The box plot of Opportunity Duration (Days) shows wide variation:
  - Most “Not Completed” learners still have active or very long opportunity durations, indicating unfinished or ongoing tasks.
  - “Completed” cases (few in the dataset) cluster within a reasonable duration range, while outliers suggest unusually long or short completions.
- Outliers may represent data entry issues or unique learners finishing abnormally early/late.

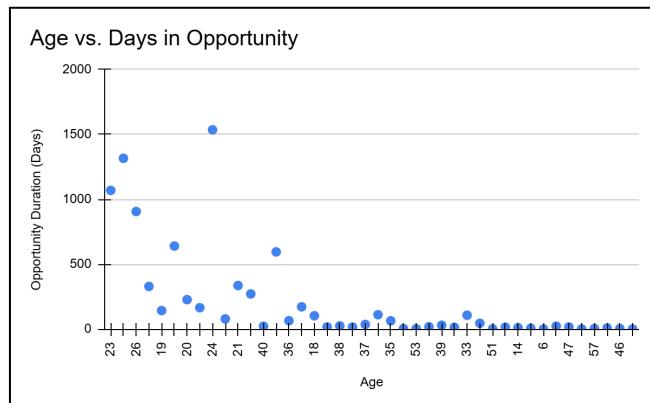


## Patterns and Correlations

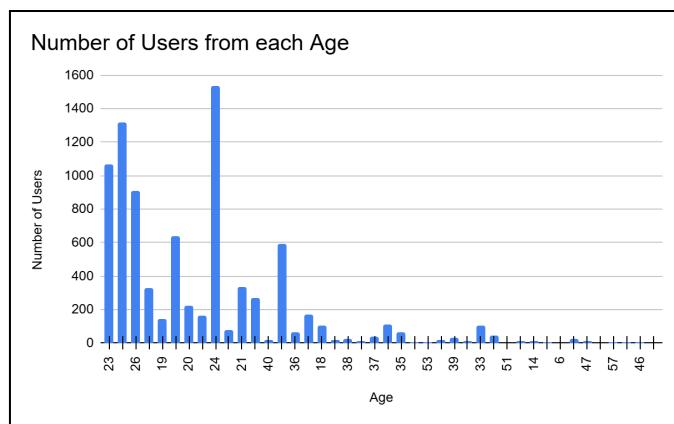
- This chart shows the Status of Students who have all applied for an opportunity from Excelerate. There are many different outcomes from applying as seen in the chart below. This gives viewers an idea of roughly how many students finish the opportunity they sign up for.



- A comparison can be made between the age of the student and how many the duration of their opportunity in days. This comparison shows the correlations between how much time a student has available based on their age.



- Next is statistics of the ages of users. As seen in the Bar Chart below, 24 seems to be the most popular age of students in opportunities at Excelerate meaning the majority audience is most likely finishing a Master's degree in college, or has completely finished college.



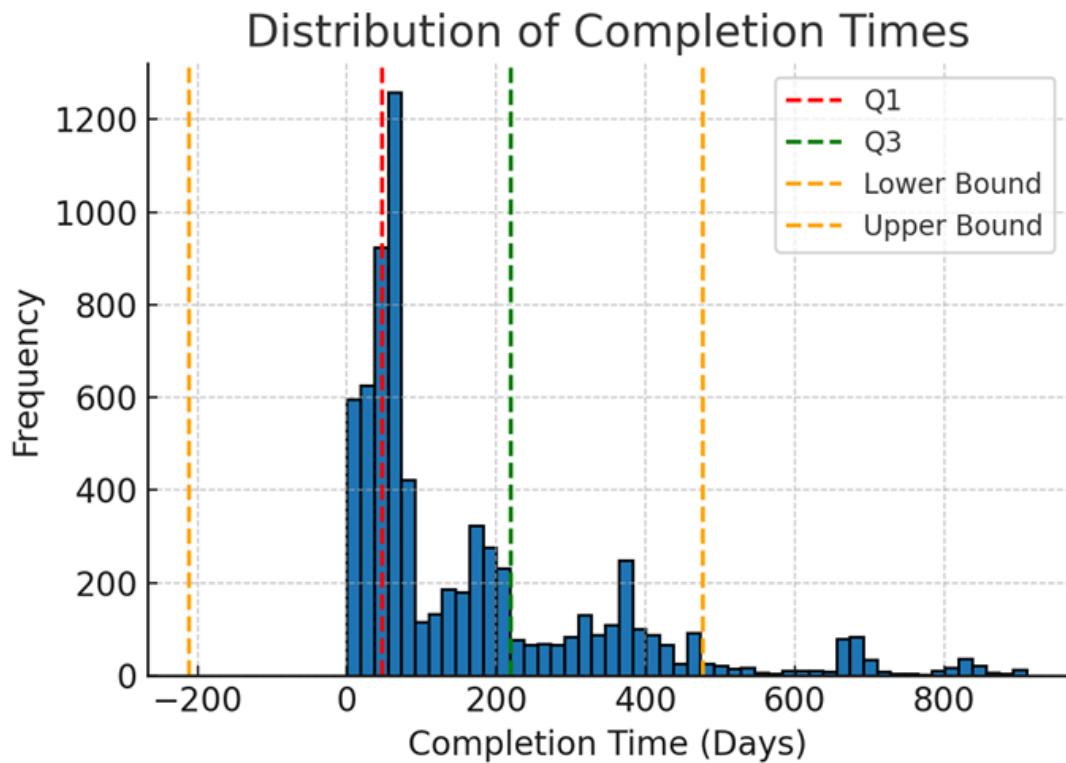
## Outliers and Anomalies

## Completion Rate Outliers

### Status Distribution:

- Rejected: 3,443 (41.7%)
- Team Allocated: 3,162 (38.3%)
- Started: 748 (9.1%)
- Dropped Out: 596 (7.2%)

The 0.4% completion rate indicates fundamental program failure. The large gap between team allocation (38.3%) and program start (9.1%) suggests critical engagement issues.



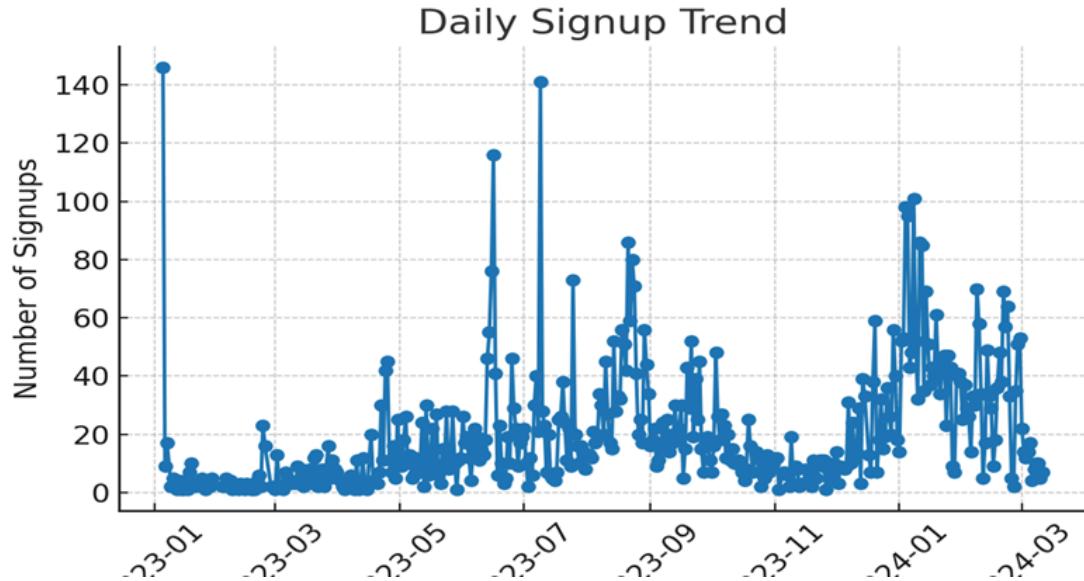
### Completion Time Statistics:

- Median Duration: 285 days (9.5 months)
- Q1-Q3 Range: 125-841 days
- Very Long Programs: 5,637 learners (68.2%) take >1 year
- Duration Outliers: 3,654 (44.2%) outside normal bounds

Programs taking 2-3 years to complete are unsustainable and contribute to high dropout rates.

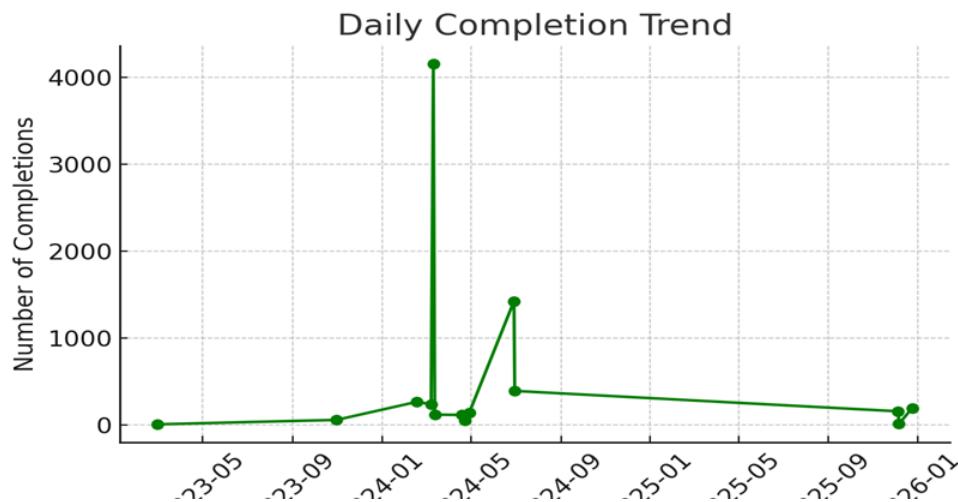
## Low Completion Days

- Mid-January Period: Consistent drops during January 12-17, 2023
- Holiday Transition Periods: Significant decreases following major holidays
- Month-End Periods: Regular drops around 26th-30th of months
- Weekend Patterns: Saturday-Sunday represent only 23% of total signups



#### **Completion Trend Issues:**

- Seasonal Completion Drops: March 2023 showed near-zero completions
- Future Date Anomalies: December 2025 completion dates suggest data quality issues
- Irregular Completion Spikes: Massive completion spike in early 2024 requires investigation



## **Recommendations**

## **Target Peak Days**

To increase enrollment efficiency, concentrate marketing efforts on days with high signup rates. According to analysis, Thursday has the highest signup volume (19.1%, or 1,576 signups), followed by Friday (16.4%, or 1,352 signups).

Plan major program launches for Thursday releases and allocate 70% of your marketing budget to Thursday-Friday campaigns. Late afternoon is the best time to deliver the campaign because hourly performance peaks at 4:00 PM (561 signups) and 6:00 PM (504 signups). Since Saturday and Sunday together only make up 23% of all signups, weekend marketing should be drastically cut. This is an inefficient use of resources.

## **Investigate Drops**

Given that 43.2% of users apply on the same day they sign up, the critical 41.7% rejection rate that affects 3,443 applicants necessitates immediate investigation. This suggests a high level of initial interest followed by a pattern of unsuccessful applications. Another crucial drop point that needs immediate attention is the 29% engagement gap between team allocation (38.3%) and program start (9.1%).

In order to identify obstacles and carry out emergency activation campaigns, survey the 3,162 team-allocated learners who have not yet begun programs. Examine the requirements and procedures for applications to learn why almost half of prospective candidates are turned down, particularly considering that the median application time is only three days.

## **Support Long-Tail Users**

Address the duration crisis, which occurs when 68.2% of students (5,637) take more than a year to finish programs, with a median completion time of 285 days.

Establish required milestone reviews every 90 days and conduct weekly check-ins for all students who have completed more than 180 days. Break existing long programs into 30-60 day modules to prevent learner fatigue and maintain engagement.

Personalised acceleration plans and more resources are required for the 44.2% of students who were identified as duration outliers. To avoid indefinite program occupancy, set completion deadlines and renewal requirements for programs that last longer than 365 days.

## **Segment Strategies**

Since nearly 80% of learners are from the United States (45.6%) and India (34.0%), localised engagement should emphasise themes of family, community, and career advancement through Eastern Time Zone campaigns for U.S. participants and IST-aligned schedules for Indian learners. Adapted strategies can be advantageous for smaller markets, such as Nigeria (8.9%).

Age-wise, gamified, mobile-friendly programs with social sharing and peer networking are most effective for the 20–25 age group (59.8%), whereas the 26–30 age group (26.8%) needs a focus on professional development, flexible evening schedules, and ROI-driven messaging. Age-based customisation combined with geographic targeting guarantees customised engagement for important learner groups.

## Conclusion

### Key Insights

- Signups showed steady growth with clear peak times — especially in August 2023 and January 2024. These periods likely connect to academic calendars and promotion efforts, showing when learners are most active.
- Completions did not keep up with signups. Most learners signed up but did not finish, which points to an engagement problem.
- Completion times also varied a lot, with some learners taking much longer than expected, suggesting that extra support or follow-ups may be needed.

### Next Steps

- Look deeper into which groups of learners (by age, country, or institution) are more or less likely to complete.
- Study whether factors like opportunity type or duration affect completion rates.
- Try simple AI-driven ideas like reminder messages or progress tracking to help boost completion.
- Use these insights to plan better outreach and support so that high signup numbers also turn into high completion rates.